Estimating the Attributable Disease Burden and Effects of Interhospital Patient Sharing on *Clostridium difficile* Infections

Daniel K. Sewell PhD, Jacob E. Simmering PhD, Samuel Justice MS, Sriram V. Pemmaraju PhD, Alberto M. Segre PhD and Philip M. Polgreen MD

1Department of Biostatistics, University of Iowa, Iowa City, Iowa, 2Department of Computer Science, University of Iowa, Iowa City, Iowa, 3Department of Statistics, University of Iowa, Iowa City, Iowa and 4Department of Internal Medicine, University of Iowa, Iowa City, Iowa for the CDC MInD-Healthcare Program

**Abstract**

Objective: To estimate the burden of *Clostridium difficile* infections (CDIs) due to interfacility patient sharing at regional and hospital levels.

Design: Retrospective observational study.

Methods: We used data from the Healthcare Cost and Utilization Project California State Inpatient Database (2005–2011) to identify 26,878,498 admissions and 532,925 patient transfers. We constructed a weighted, directed network among the hospitals by defining an edge between 2 hospitals to be the monthly average number of patients discharged from one hospital and admitted to another on the same day. We then used a network autocorrelation model to study the effect of the patient-sharing network on the monthly average number of CDI cases per hospital, and we estimated the proportion of CDI cases attributable to the network.

Results: We found that 13% (95% confidence interval [CI], 7.6%–18%) of CDI cases were due to diffusion through the patient-sharing network. The network autocorrelation parameter was estimated at 5.0 (95% CI, 3.0–6.9). An increase in the number of patients transferred into and/or an increased CDI rate at the hospitals from which those patients originated led to an increase in the number of CDIs in the receiving hospital.

Conclusions: A minority but substantial burden of CDI infections are attributable to hospital transfers. A hospital’s infection control may thus be nontrivially influenced by its neighboring hospitals. This work adds to the growing body of evidence that intervention strategies designed to minimize HAIs should be done at the regional rather than local level.

(Received 27 August 2018; accepted 7 March 2019)
Methods

Data

The study results were based on the State Inpatient Database (SID), an administrative claims database available through the Healthcare Utilization Project. The SID contains information from discharge abstracts for all inpatient visits to a nonfederal hospital for specific states and years. In this analysis, we used California data for the period from January 2005 to November 2011. This dataset consisted of 27,200,873 records from 408 hospitals. We excluded from the analysis those hospitals that did not appear over the entirety of the study period. Thus, we retained 385 hospitals, which accounted for >98.8% of the patient records. We focused on California based on its large size and because its geography minimizes transfers to and from other states.

Critically, this dataset contained data regarding return visits, which allowed us to track patients over multiple visits. Thus, we were able to discern transfers between hospitals by considering patients who had common discharge and admission dates involving 2 distinct hospitals. We defined a patient transfer as a patient discharged from one hospital (the source hospital, $i$) and then admitted to another hospital (the target hospital, $t$) on the same day. We defined a weighted network as a set of nodes and a set of pairs of nodes, for which each pair has an assigned value. We constructed such a weighted network in which the nodes are the 385 hospitals and a weighted edge between hospital $s$ and hospital $t$ is set to be equal to the monthly average number of transfers from $s$ to $t$.

CDI cases were identified as inpatient visits with an International Classification of Disease, Ninth Revision, Clinical Modification (ICD-9-CM) diagnosis code of 008.45. Principal CDI diagnoses were excluded because these were unlikely to be related to HAIs. Both the sensitivity and specificity of this code are high (78% and 99.7% respectively). Our primary outcome of interest was the monthly average number of CDI cases at each hospital. We also computed the monthly average number of admissions, median length of stay, median number of diagnoses (ie, the number of diagnosis codes listed for the patient at the time of discharge) per inpatient, and percent of inpatients aged 65 years and older.

Statistical analysis

Our goal was to determine the expected increase in monthly average number of CDI cases due to the patient transfer network acting as a vector for disease diffusion. To account for the varying patient populations associated with the 385 hospitals, we considered as covariates the median length of stay, the median number of diagnoses per inpatient visit, and the percent of patients aged over 65. To account for potential community effects, we included tensor product splines using the latitude/longitude coordinates of the hospitals, and we used Akaike information criterion (AIC) to select the number of knots. Additionally, we included the hospital’s monthly average admission count as a covariate. If the expected number of CDI cases was simply a proportion of a hospital’s total admissions, this would be the only significant covariate. To answer our primary research questions, we also incorporated the transfer network contagion effect into the modeling framework defined to be the change in the number of CDI cases due to patient transfers.

The intuition leading to our statistical modelling choice is as follows: Assume that the probability of a patient transferring out of a hospital being symptomatically or asymptotically infected or contaminated with Clostridium difficile is proportional to that hospital’s CDI rate. The expected number of contagious patients being transferred out of a source hospital and into a target hospital is then proportional to the product of the total number of transfers from the source to the target and the source’s CDI rate. We can then investigate the effect on the target hospital’s CDI cases from the total expected number of contagious transfer patients. The model in terms of the $i$th hospital can be informally conceptualized as

\[
\text{NAM}(\text{CDI}_i) = \text{NAM}(\text{Admissions})_i + \text{Median \# LOS} + \text{Median \# Dx} + \text{(% Over 65\%)} + \sum_{j\neq i} (\text{# transfers from } j \text{ to } i) \times (\text{Avg CDI Rate})_i.
\]

Statistically valid inference can be achieved by implementing a network autocorrelation model (NAM) and finding the corresponding maximum likelihood estimators and their sampling distributions. The NAM estimates the effects of covariates as well as the network on the mean of the response variable. This framework appropriately accounts for the dependencies between the observations induced by the connectivity through patient sharing, thereby yielding consistent estimators and correct standard errors (see the Supplemental Material online).

By allowing us to estimate the network effect on the mean monthly average number of CDI cases, the NAM also provided an opportunity to evaluate a counterfactual framework. Specifically, we computed the expected total number of monthly CDI cases accounting for both the covariates and the network contagion effect. We then estimated the expected total number of monthly CDI cases as if we were capable of negating the network contagion effect while keeping all else constant. We then considered the percent decrease in expected total number of monthly CDI cases due to this negation. In this way, we were able to assess the percent of CDI cases affected by the patient sharing network (see the Supplementary Material online for details of the counterfactual framework).

To ensure that the network term in the NAM was in fact estimating a contagion effect rather than acting as a proxy for in-degree (the number of other hospitals transferring patients to a given hospital), which could itself be a proxy for any number of confounding factors such as available services, we ran the model with in-degree included as a covariate to verify and reassess the statistical significance of the network term in the NAM. Because this term remained significant we were confident that the network term in the NAM had captured the desired contagion effect. However, the interpretability of the model decreases with both in-degree and the network term included in the model; thus, it was more useful to exclude in-degree as a covariate as it was not statistically significant.

Results

We identified 532,925 transfers over the 7-year period. Of all 147,840 possible directed pairs of the 385 hospitals, 14.4% shared at least 1 patient. Of those directed pairs of hospitals that did share patients, the mean (SD) of the monthly average number of transfers was 0.302 (SD, 1.37; range, 0.0120–64.4).

Figure 1 shows the 385 hospitals included in the study, and a line connects any pair of hospitals that transferred >1 patient per year from one to the other over the observed time frame. Patient sharing is apparently highly pervasive and not constrained by geographic considerations. Table 1 provides the descriptive statistics for the response variable and the covariates in our NAM (see the Supplementary Material online for more detail).
Table 2 provides the results from the NAM. Increasing the monthly average total admissions by 100 led to an expected increase in monthly average number of CDI cases by 0.59 (95% CI, 0.54–0.63). This result reflects the overall proportion of admissions with a secondary diagnosis of CDI (0.60%). Increasing the median number of diagnoses per patient by 1 led to an expected increase in monthly average number of CDI of 0.51 (95% CI, 0.40–0.68). After comparing 64 combinations of latitude and longitude cubic tensor product splines ranging from 3 to 10 marginal degrees of freedom each, the AIC chose 3 degrees of freedom for both. The estimated effect on the monthly average number of CDIs due to the spatial effect is shown in Figure 1. The likelihood ratio test, however, yielded a P value of 0.70, indicating that the spatial component was not statistically significant.

The final 2 rows of Table 2 correspond to the patient transfers. The network contagion effect was estimated to be 5.0 (95% CI, 3.0–6.9), indicating that an increase in the number of patients transferred into and/or an increased CDI rate at the hospitals from which those patients originated led to an increase in the number of CDIs in the receiving hospital. This effect may be interpreted as follows. A source hospital with a CDI rate of r will increase the expected number of CDI cases in a target hospital by 5 for every 1/r patients transferred from source to target. For example, if a source hospital has a CDI rate of 0.01, every hundred transfers from this hospital to another will increase the expected number of CDI cases in the target hospital by 5. Figure 2 illustrates this effect by showing the expected increase in the number of CDI cases due to patients transferring out of a source hospital as a function of both the number of transfers out and the CDI rate of the source. The white points provide the observed values for the 385 hospitals. The final row of Table 2 corresponds to our counterfactual framework. If the patient transfers were eliminated, our model predicts a 13% (95% CI, 7.6%–18%) decrease in the monthly number of CDI cases.

### Table 1. Descriptive Statistics for the Hospital-Level Variables in the Network Autocorrelation Model (NAM) Corresponding to 385 Hospitals From 2005 to 2011

<table>
<thead>
<tr>
<th>Variable</th>
<th>Median</th>
<th>IQR</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monthly avg. no. of CDI cases</td>
<td>3.65</td>
<td>6.94</td>
<td>0.00</td>
<td>52.3</td>
</tr>
<tr>
<td>Monthly avg. total admissions</td>
<td>601</td>
<td>1,070</td>
<td>3.27</td>
<td>4,560</td>
</tr>
<tr>
<td>Median LOS, d</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>42.5</td>
</tr>
<tr>
<td>Median no. of diagnoses</td>
<td>6</td>
<td>2</td>
<td>1</td>
<td>22</td>
</tr>
<tr>
<td>Patients aged &gt;65 y, %</td>
<td>33.0</td>
<td>21.1</td>
<td>0.00</td>
<td>86.5</td>
</tr>
</tbody>
</table>

Note. IQR, interquartile range; CDI, Clostridium difficile infection; LOS, length of stay.

### Table 2. Results from the Network Autocorrelation Model (NAM)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>95% CI</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>−4.44</td>
<td>(−6.44 to −2.44)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Total admissions</td>
<td>0.00587</td>
<td>(0.00543–0.00631)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Median LOS, d</td>
<td>−0.0267</td>
<td>(−0.0952 to 0.0417)</td>
<td>.444</td>
</tr>
<tr>
<td>Median no. of diagnoses</td>
<td>0.514</td>
<td>(0.350–0.679)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Patients aged &gt;65 y, %</td>
<td>0.0146</td>
<td>(0.00536–0.0345)</td>
<td>.152</td>
</tr>
<tr>
<td>Geospatial effect</td>
<td>. . .</td>
<td>. . .</td>
<td>. . .</td>
</tr>
<tr>
<td>Network contagion effect</td>
<td>4.99</td>
<td>(3.03–6.94)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>% attributable to patient sharing</td>
<td>12.6</td>
<td>(7.60–17.7)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

Note. CI, confidence interval; LOS, length of stay; CDI, Clostridium difficile infection.

*Parameter determining the number of CDI cases exported by a source hospital based on the source’s contamination level and the number of patients being transferred out of the source. See text.

*Quantity from counterfactual framework describing the percent of all CDI cases attributable to patient transfers. See text.
Discussion

We have demonstrated a statistical method of estimating both the effect of patient transfers on disease rates and the proportion of cases due to the network. Our results demonstrate that CDI rates are associated with hospital transfers after controlling for hospital-level factors and community effects. Across California, if the contamination effect of hospital transfers were eliminated, our estimates would indicate a statewide decrease in CDI of 13%, corresponding to a reduction in 247 expected statewide number of cases each month. At a hospital level, we were able to estimate the increase in the expected number of CDI cases due to transfers as a function of the CDI rate of the source hospital and the number of patient transfers.

In addition to our transfer-related results, we also found that higher rates of CDI were associated with a greater number of admissions. Hospitals with patients with more comorbidities, measured by their number of diagnoses, also had higher rates of CDI. All of these hospital-level risk factors confirm previous findings at the individual level.8,14–18,27

Our method of estimating the burden of disease attributable to hospital transfers used a statistical model and a counterfactual-like approach. This approach was based on conditional expectations; we compared the expected number of CDI cases given that there is a network effect versus the expected number of CDI cases given that the network effect is negated. Our results indicate that hospital transfers had a significant effect on predicted CDI rates. Specifically, we have demonstrated that if the effect of hospital transfers were mitigated by some process of inducing immunity (eg, by increased screening or vaccination), predicted CDI rates would fall. Thus, our results allowed us to estimate, or at least place an upper bound on, the burden of CDI attributable to patient transfers. In prior work, controlling for a variety of hospital-level risk factors, Simmering et al30 found that an increase of 1 in log in-degree was associated with an increase of 4.8% in the CDI rate. Our new results corroborate these prior findings while providing (1) more confidence that our network term was not acting as a proxy for some confounding factor, and (2) hospital-level burden estimates.

Travel is an important driver of the spread of a broad range of infectious diseases, including pathogens associated with spread within healthcare facilities.43 Thus, it is not surprising that hospital transfers, a specialized case of travel, have been implicated in the spread of pathogens. Hospital transfers have been implicated in the spread of rapidly emerging infections (eg, severe acute respiratory syndrome [SARS]) and have been associated with the spread of MRSA and MDROs in many different geographic regions.15,31,32,44–47 Transfers of patients demonstrate the interconnectedness of healthcare systems. Accordingly, efforts to control the spread of infections at one hospital transfers may, to some extent, be a function of higher levels of frailty and higher levels of comorbidities among patients being transferred from hospitals with higher CDI rates. Alternatively, the hospital-transfer structure and its association with risk may be measured changes in antimicrobial exposure in different hospital environments. Indeed, patients transferred for medical reasons may not only have greater severity of illness but may also be more likely to be exposed to broad-spectrum antibiotics. Nevertheless, our statistical framework can be thought of as an upper bound of the infections attributable to patient transfers. In our specific example, we would expect a significant but overall modest reduction of CDI if we could eliminate transfers altogether. Thus, our framework can be used to estimate the likely impact of population- or community-based interventions involving reducing risk of healthcare-associated infections among transfer patients. Although we can use the same framework with different hospital-associated infections, we would, of course, expect the results to change from organism to organism depending upon transmissibility.

Our study has several limitations. First, we used administrative data to determine both outcome and predictor variables. We did not have access to microbiologic testing data or the type of test used to diagnose CDI within our dataset. Although there are other sources of CDI infection data (eg, data reported directly to public health agencies), they do not contain the appropriate transfer-network data nor do they contain information on patients that did not get CDI. Second, our data did not contain antibiotic prescribing data at an individual level nor did we have antibiotic prescribing data in aggregate at the receiving- or transfer-hospital level. Antimicrobial prescribing at both individual and population levels are known to be important drivers of CDI.3,4,12,29 Third, we may be missing some CDI cases attributable to hospitals in cases in which CDI presented after hospital discharge.49 Fourth, our data do not provide information on transfers originating from another state; hence, some patient transfers may have been missing. Fifth, although we have tried to account for patient population characteristics and community effects, other sources of CDI that could impact our conclusions may not have been unaccounted for. Sixth, we only used data from 1 state, and our results need to be replicated in other states to improve their generalizability. Future work with other states and including other variables is needed to increase the generalizability and the reliability of our estimates.

Despite these limitations, we have clearly demonstrated that hospital transfers are associated with higher CDI rates and that reducing transfers may decrease the number of CDI cases at a population level. In addition, our methodological approach provides a framework for estimating the burden of particular hospital-acquired infections that are attributable to hospital transfers for a wide range of pathogens. Thus, this modeling framework could be useful in estimating the impact of transfer-focused infection control interventions.

**Supplementary material.** To view supplementary material for this article, please visit https://doi.org/10.1017/ice.2019.73.

**Author ORCIDs.** Daniel K. Sewell, 0000-0002-9238-4026

**Financial support.** This work was supported by the US Centers for Disease Control and Prevention (grant no. 5 U01 CK000531-02).